

SEQUENCE LISTING

5 (1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: CLOUGH et al

10 (B) STREET: National Institute for Medical Research, Mill
Hill

(C) CITY: London

(E) COUNTRY: United Kingdom

(F) POSTAL CODE (ZIP): NW7 1AA

15 (ii) TITLE OF INVENTION: AN EF-TU PROTEIN ENCODED ON THE PLASTID DNA
OF THE MALARIA PARASITE AND PROTEIN SYNTHESIS INHIBITORS
EFFECTIVE AS ANTI-MALARIAL COMPOUNDS

20 (iii) NUMBER OF SEQUENCES: 2

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-DOS/MS-DOS

25 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(v) CURRENT APPLICATION DATA:

APPLICATION NUMBER: US NA

30 (2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1230 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

10

(A) NAME/KEY: CDS
 (B) LOCATION:1..1230

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

15

ATG AAT AAT AAA TTA TTT TTA AGA AAT AAA CAA CAT ATA AAT TTA GGT	48
Met Asn Asn Lys Leu Phe Leu Arg Asn Lys Gln His Ile Asn Leu Gly	
1 5 10 15	

20

ACT ATA GGG CAT GTA GAT CAT GGA AAA ACT ACA TTA ACA ACA GCT ATA	96
Thr Ile Gly His Val Asp His Gly Lys Thr Thr Leu Thr Thr Ala Ile	
20 25 30	

25

TCT TAT TTA TTA AAT TTA CAA GGA TTA TCA AAA AAA TAT AAT TAT TCA	144
Ser Tyr Leu Leu Asn Leu Gln Gly Leu Ser Lys Lys Tyr Asn Tyr Ser	
35 40 45	

30

GAT ATT GAT TCA GCT CCA GAA GAA AAA ATA AGA GGT ATT ACA ATA AAT	192
Asp Ile Asp Ser Ala Pro Glu Glu Lys Ile Arg Gly Ile Thr Ile Asn	
50 55 60	

ACA ACA CAT ATT GAA TAT GAA ACT TTA ACA AAA CAT TGT GCT CAT ATA	240
---	-----

	Thr Thr His Ile Glu Tyr Glu Thr Leu Thr Lys His Cys Ala His Ile	
	65 70 75 80	
5	GAT TGT CCA GGA CAT TCC GAT TAT ATT AAA AAT ATG ATT ATA GGA GCC Asp Cys Pro Gly His Ser Asp Tyr Ile Lys Asn Met Ile Ile Gly Ala	288
	85 90 95	
10	ACA CAA ATG GAT ATA GCA ATT TTA GTA ATA TCT ATA ATA GAT GGT ATA Thr Gln Met Asp Ile Ala Ile Leu Val Ile Ser Ile Ile Asp Gly Ile	336
	100 105 110	
15	ATG CCT CAA ACT TAT GAA CAT TTA TTA TTA ATA AAA CAA ATA GGT ATA Met Pro Gln Thr Tyr Glu His Leu Leu Leu Ile Lys Gln Ile Gly Ile	384
	115 120 125	
20	AAA AAT ATA ATT ATT TTT TTA AAT AAA GAA GAT TTA TGT GAT GAT GTT Lys Asn Ile Ile Ile Phe Leu Asn Lys Glu Asp Leu Cys Asp Asp Val	432
	130 135 140	
25	GAA TTA ATA GAT TTT ATA AAA TTA GAA GTA AAT GAA TTA TTA ATT AAA Glu Leu Ile Asp Phe Ile Lys Leu Glu Val Asn Glu Leu Leu Ile Lys	480
	145 150 155 160	
30	TAT AAT TTT GAT TTA AAT TAT ATA CAT ATA TTA ACT GGT TCA GCA TTA Tyr Asn Phe Asp Leu Asn Tyr Ile His Ile Leu Thr Gly Ser Ala Leu	528
	165 170 175	
35	AAT GTA ATA AAT ATA ATT CAA AAA AAT AAG GAT TAT GAA TTA ATA AAA Asn Val Ile Asn Ile Ile Gln Lys Asn Lys Asp Tyr Glu Leu Ile Lys	576
	180 185 190	
40	TCT AAT ATT TGG ATA CAA AAA TTA AAT AAT TTA ATT CAA ATA ATT GAT	624

Ser Asn Ile Trp Ile Gln Lys Leu Asn Asn Leu Ile Gln Ile Ile Asp
 195 200 205

5 AAT ATT ATA ATA CCT ACT AGA AAA ATT AAT GAT TAC TTT TTA ATG TCA 672
 Asn Ile Ile Ile Pro Thr Arg Lys Ile Asn Asp Tyr Phe Leu Met Ser
 210 215 220

10 ATA GAA GAT GTA TTT TCT ATA ACA GGT AGA GGT ACA GTA GTA ACA GGT 720
 Ile Glu Asp Val Phe Ser Ile Thr Gly Arg Gly Thr Val Val Thr Gly
 225 230 235 240

15 AAG ATT GAA CAA GGA TGT ATA AAT TTA AAT GAT GAA ATT GAA ATT TTA 768
 Lys Ile Glu Gln Gly Cys Ile Asn Leu Asn Asp Glu Ile Glu Ile Leu
 245 250 255

AAA TTT GAA AAA TCA TCT CCT AAT TTA ACA ACA GTT ATA GGA TTA GAA 816
 Lys Phe Glu Lys Ser Ser Pro Asn Leu Thr Thr Val Ile Gly Leu Glu
 260 265 270

20 ATG TTT AAA AAA CAA TTA ACA CAA GCA CAA TCC GGA GAT AAT GTA GGT 864
 Met Phe Lys Lys Gln Leu Thr Gln Ala Gln Ser Gly Asp Asn Val Gly
 275 280 285

25 ATT TTA TTA AGA AAT ATT CAA AAA AAA GAT ATA AAA AGA GGT ATG ATT 912
 Ile Leu Leu Arg Asn Ile Gln Lys Lys Asp Ile Lys Arg Gly Met Ile
 290 295 300

30 TTA GCA ACA CCT AAT AAA TTA AAA GTA TAT AAG TCT TTT ATA GCT GAA 960
 Leu Ala Thr Pro Asn Lys Leu Lys Val Tyr Lys Ser Phe Ile Ala Glu
 305 310 315 320

ACA TAT ATT TTA ACT AAA GAA GAA GGT GGT CGT CAT AAA CCT TTT AAT 1008

Thr Tyr Ile Leu Thr Lys Glu Glu Gly Gly Arg His Lys Pro Phe Asn
 325 330 335

ATT GGA TAT AAA CCT CAA TTT TTT ATT CGT ACA GTA GAT GTT ACT GGA 1056
 5 Ile Gly Tyr Lys Pro Gln Phe Phe Ile Arg Thr Val Asp Val Thr Gly
 340 345 350

GAA ATT AAA AAT ATA TAT TTA AAT GAA AAT GTA CAA AAA GTA GCT ATA 1104
 10 Glu Ile Lys Asn Ile Tyr Leu Asn Glu Asn Val Gln Lys Val Ala Ile
 355 360 365

CCT GGA GAT AAA ATA ACA TTA CAT ATT GAA TTA AAA CAT TAT ATA GTG 1152
 15 Pro Gly Asp Lys Ile Thr Leu His Ile Glu Leu Lys His Tyr Ile Val
 370 375 380

TTG ACA TTA AAT ATG AAA TTT TCT ATT AGA GAA GGA GGA AAA ACA ATA 1200
 Leu Thr Leu Asn Met Lys Phe Ser Ile Arg Glu Gly Gly Lys Thr Ile
 385 390 395 400

GGA GCA GGT ATT ATA ACA GAA ATA AAA AAT 1230
 20 Gly Ala Gly Ile Ile Thr Glu Ile Lys Asn
 405 410

25 (2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 410 amino acids

(B) TYPE: amino acid

30 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Asn Asn Lys Leu Phe Leu Arg Asn Lys Gln His Ile Asn Leu Gly
1 5 10 15

5
Thr Ile Gly His Val Asp His Gly Lys Thr Thr Leu Thr Thr Ala Ile
20 25 30

Ser Tyr Leu Leu Asn Leu Gln Gly Leu Ser Lys Lys Tyr Asn Tyr Ser
10 35 40 45

Asp Ile Asp Ser Ala Pro Glu Glu Lys Ile Arg Gly Ile Thr Ile Asn
50 55 60

15
Thr Thr His Ile Glu Tyr Glu Thr Leu Thr Lys His Cys Ala His Ile
65 70 75 80

Asp Cys Pro Gly His Ser Asp Tyr Ile Lys Asn Met Ile Ile Gly Ala
85 90 95

20
Thr Gln Met Asp Ile Ala Ile Leu Val Ile Ser Ile Ile Asp Gly Ile
100 105 110

Met Pro Gln Thr Tyr Glu His Leu Leu Leu Ile Lys Gln Ile Gly Ile
25 115 120 125

Lys Asn Ile Ile Ile Phe Leu Asn Lys Glu Asp Leu Cys Asp Asp Val
130 135 140

30
Glu Leu Ile Asp Phe Ile Lys Leu Glu Val Asn Glu Leu Leu Ile Lys
145 150 155 160

Tyr Asn Phe Asp Leu Asn Tyr Ile His Ile Leu Thr Gly Ser Ala Leu
165 170 175

5 Asn Val Ile Asn Ile Ile Gln Lys Asn Lys Asp Tyr Glu Leu Ile Lys
180 185 190

Ser Asn Ile Trp Ile Gln Lys Leu Asn Asn Leu Ile Gln Ile Ile Asp
195 200 205

10 Asn Ile Ile Ile Pro Thr Arg Lys Ile Asn Asp Tyr Phe Leu Met Ser
210 215 220

Ile Glu Asp Val Phe Ser Ile Thr Gly Arg Gly Thr Val Val Thr Gly
225 230 235 240

15 Lys Ile Glu Gln Gly Cys Ile Asn Leu Asn Asp Glu Ile Glu Ile Leu
245 250 255

20 Lys Phe Glu Lys Ser Ser Pro Asn Leu Thr Thr Val Ile Gly Leu Glu
260 265 270

Met Phe Lys Lys Gln Leu Thr Gln Ala Gln Ser Gly Asp Asn Val Gly
275 280 285

25 Ile Leu Leu Arg Asn Ile Gln Lys Lys Asp Ile Lys Arg Gly Met Ile
290 295 300

Leu Ala Thr Pro Asn Lys Leu Lys Val Tyr Lys Ser Phe Ile Ala Glu
305 310 315 320

30 Thr Tyr Ile Leu Thr Lys Glu Glu Gly Gly Arg His Lys Pro Phe Asn
325 330 335

Ile Gly Tyr Lys Pro Gln Phe Phe Ile Arg Thr Val Asp Val Thr Gly
340 345 350

5 Glu Ile Lys Asn Ile Tyr Leu Asn Glu Asn Val Gln Lys Val Ala Ile
355 360 365

Pro Gly Asp Lys Ile Thr Leu His Ile Glu Leu Lys His Tyr Ile Val
370 375 380

10 Leu Thr Leu Asn Met Lys Phe Ser Ile Arg Glu Gly Gly Lys Thr Ile
385 390 395 400

15 Gly Ala Gly Ile Ile Thr Glu Ile Lys Asn
405 410